The PrP \textsuperscript{C} C1 fragment derived from the ovine A\textsubscript{136}R\textsubscript{154}R\textsubscript{171} \textit{PRNP} allele is highly abundant in sheep brain and inhibits fibrillisation of full-length PrP\textsuperscript{C} protein \textit{in vitro}.

Lauren Campbell\textsuperscript{a}, Andrew C. Gill\textsuperscript{a}, Gillian McGovern\textsuperscript{b}, Clara M. O. Jalland\textsuperscript{c}, John Hopkins\textsuperscript{a}, Michael A. Tranulis\textsuperscript{c}, Nora Hunter\textsuperscript{a}, Wilfred Goldmann\textsuperscript{a*}

\textsuperscript{a}The Roslin Institute and Royal (Dick) School of Veterinary Studies, University of Edinburgh, Easter Bush, Midlothian, Scotland, U.K.
\textsuperscript{b}Animal Health and Veterinary Laboratories Agency (AHVLA), Pentlands Science Park, Bush Loan, Midlothian, U.K.
\textsuperscript{c}Institute of Basic Sciences and Aquatic Medicine, Department of Biochemistry and Physiology, Norwegian School of Veterinary Science, 0033 Oslo, Norway

*Corresponding author:
Correspondence to: Wilfred Goldmann, Neurobiology Division, The Roslin Institute and Royal (Dick) School of Veterinary Studies, University of Edinburgh, Easter Bush, Midlothian, EH25 9RG, U.K. Tel: 0044(0131)6519100; FAX: 0044(0131)6519105; E-mail: Wilfred.goldmann@roslin.ed.ac.uk

Email addresses: Lauren.campbell@roslin.ed.ac.uk
Andy.gill@roslin.ed.ac.uk
Gillian.mcgovern@ahvla.gsi.gov.uk
ClaraMaria.Jalland@nvh.no
John.Hopkins@roslin.ed.ac.uk
Michael.tranulis@nvh.no
Nora.hunter@roslin.ed.ac.uk
Wilfred.goldmann@roslin.ed.ac.uk
Abstract
Expression of the cellular prion protein (PrP<sup>C</sup>) is crucial for the development of prion diseases. Resistance to prion diseases can result from reduced availability of the prion protein or from amino acid changes in the prion protein sequence. We propose here that increased production of a natural PrP α-cleavage fragment, C1, is also associated with resistance to disease. We show, in brain tissue, that ARR homozygous sheep, associated with resistance to disease, produced PrP<sup>C</sup> comprised of 25% more C1 fragment than PrP<sup>C</sup> from the disease-susceptible ARQ homozygous and highly susceptible VRQ homozygous animals. Only the C1 fragment derived from the ARR allele inhibits in-vitro fibrillisation of other allelic PrP<sup>C</sup> variants. We propose that the increased α-cleavage of ovine ARR PrP<sup>C</sup> contributes to a dominant negative effect of this polymorphism on disease susceptibility. Furthermore, the significant reduction in PrP<sup>C</sup> β-cleavage product C2 in sheep of the ARR/ARR genotype compared to ARQ/ARQ and VRQ/VRQ genotypes, may add to the complexity of genetic determinants of prion disease susceptibility.

Abbreviations
PrP<sup>C</sup>, cellular prion protein; CNS, central nervous system; CJD, Creutzfeldt-Jakob disease; CWD, chronic wasting disease; BSE, bovine spongiform encephalopathy; PrP<sup>Sc</sup>, disease associated misfolded prion protein; NPU, Neuropathogenesis Unit; AHVLA, Animal Health and Veterinary Laboratories Agency; PMSF, phenylmethanesulfonylfluoride; NEM, N-ethylmaleimide; PNGase F, Peptide:N-glycosidase F; TBS, Tris-buffered saline; TBST, TBS-Tween; ALP, Alkaline Phosphatase; IPTG, isopropylthio-β-galactoside; HPLC, High pressure liquid chromatography; MES, 2-(N-Morpholino)ethanesulfonic acid sodium salt; ThT, Thioflavin T.

Keywords:
prion; transmissible spongiform encephalopathy; fibrillisation; protein-cleavage

1. Introduction
The prion protein, PrP<sup>C</sup>, is encoded by the PRNP gene and is expressed at high levels in the central nervous system (CNS) and to a lesser degree in peripheral tissues in all mammalian species. The biological function of PrP<sup>C</sup> is not clear, but roles in promoting cell survival, signal transduction and alleviating oxidative stress have been suggested [1-3]. PrP<sup>C</sup> is essential for the pathogenesis of a group of disorders called prion diseases, also known as
transmissible spongiform encephalopathies (TSEs). TSEs are fatal, neurodegenerative protein-misfolding diseases caused by unconventional infectious agents, often referred to as prions [4,5]. Naturally occurring prion diseases are Creutzfeldt-Jakob disease (CJD) in man, chronic wasting disease (CWD) in deer, bovine spongiform encephalopathy (BSE) in cattle and scrapie in sheep. A marker of prion disease infection is the slow accumulation of a protease-resistant isoform of the prion protein, designated PrPSc, in the CNS and the peripheral lymphoid system [6]. PrPSc is converted from PrPC by a seeded polymerisation mechanism leading to a range of morphologically different aggregates, ranging from fibrils to diffuse amyloid plaques [7]. The process can be mimicked in vitro, whereby recombinant PrPC can be induced to transit from a soluble alpha-helical state into insoluble, highly ordered amyloid fibrils by a nucleation-dependent mechanism [8]. Most mammalian species exhibit PRNP gene polymorphisms which encode PrPC protein sequence variants and several of these have been shown to modulate susceptibility, incubation period or pathology of prion diseases [9]. One of the best described genetic associations between PRNP gene polymorphisms (PrPC variants) and prion diseases applies to sheep scrapie. Although the genetic determinants of susceptibility to scrapie infection and incubation period length in sheep are highly complex, involving more than ten polymorphic positions, three substitutions of major importance are in codons 136, 154 and 171. Ovine PrPC variants are therefore described by the amino acid expressed at these codons, such as A136R154R171 (ARR), ARQ and VRQ etc. [10-12]. It is still a matter of debate exactly how PrPC amino acid substitutions cause differences in disease susceptibility and pathogenesis. The efficiency of in-vitro conversion of PrPC to PrPSc has been shown to vary for different PrPC variants [13-15] and it has been suggested that this variation is linked to differential protein stability [14], but it is likely that in vivo additional factors contribute to disease phenotype differences. Ovine PrPC is expressed in brain as di- (33-35 kDa), mono- (30-32 kDa) or un-glycosylated (27 kDa) protein in variable ratios. Cellular processing of PrPC involves two well-documented proteolytic cleavage events. For ovine PrPC, α-cleavage of the peptide-bond between His114 and Val115 [18,19] creates two polypeptides, of which C1 is the C-terminal fragment that resides alongside full length PrPC on the cell membrane [20]. C1 also appears in di- (25-27kDa) mono- (21-23kDa) and un-glycosylated (17kDa) forms [21-23]. The majority of the corresponding N-terminal, 9kDa N1 fragment is released from the cell by shedding [24]. The products of α-cleavage of PrPC have been observed in the brains of a variety of mammals with diverse susceptibility to TSE diseases [20-22, 25-27]. An alternative β-cleavage, at Gly92 of ovine PrPC, generates the fragments C2 and N2, but occurs at a much lower level in
healthy animals than α-cleavage and appears to be a response to oxidative stress [28]. It is a long held view that prion disease incubation periods are correlated with the amount of PrP\textsuperscript{C}. This has been demonstrated in various transgenic mouse models in which incubation periods of experimentally induced prion diseases are inversely correlated with the PrP\textsuperscript{C} protein expression [29, 30]. However, often in these studies the levels of full-length PrP\textsuperscript{C} and its proteolytic fragments have not been differentiated. Recently it was shown conclusively that the C1 fragment itself does not convert into a protease-resistant isoform in scrapie challenged transgenic mice expressing only C1. Furthermore, when the C1 fragment was co-expressed with full-length PrP\textsuperscript{C} incubation periods were extended [31]. It has also been shown that cell lines with naturally higher levels of α-cleavage show enhanced resistance to prion infection [32]. This raises the possibility that PrP\textsuperscript{C} cleavage may control disease by either reducing the amount of full length PrP\textsuperscript{C} available for conversion or by producing different levels of the C1 fragment, which would act as inhibitory modulators of conversion. In sheep, this control is likely to be associated with PRNP genotype. We have found that proteolytic processing of ovine PrP\textsuperscript{C} is PRNP genotype dependent, with increased amounts of the C1 fragment and decreased amounts of the C2 fragment in brain tissue from sheep associated with resistance to scrapie. Furthermore, we have shown that a recombinant protein comprising the C1 fragment derived from the ovine ARR variant has the ability to inhibit or delay fibrillisation of full length PrP, a key step in the formation of disease associated amyloid.

2. Material and methods

2.1 Brain tissue preparations

Sheep used in this study were obtained from the Roslin Institute Cheviot flock (formerly known as the NPU Cheviot flock) [33] and from a scrapie free flock (AHVLA) (both together in this paper called Roslin sheep) or from random sampling at Norwegian abattoirs (in this paper called Norwegian sheep). All sheep were arginine homozygous in codon 154 (RR\textsubscript{154}). The VRQ/VRQ, VRQ/ARQ and ARQ/ARQ genotypes are collectively referred to as QQ\textsubscript{171}, the ARR/ARR genotype as RR\textsubscript{171}. For some Roslin sheep we retained 1 cm midsectional slices from the left side of the brain, from which cores of approximately 5 mm diameter were removed from the cortex, cerebellum, medulla, thalamus, hypothalamus, mid brain and pons. Following post mortem, all tissues were immediately stored at -70 °C. Tissues were manually homogenised in lysis buffer (5 % NP-40 (v/v), 12.1 mM Sodium deoxycholate in PBS) with protease inhibitors (10 µM PMSF, 10 µM NEM or Complete Mini Tablets, Roche) to make a 10 % (w/v) homogenate. The homogenate was clarified by
centrifugation at 2000 rpm at 4 °C for 10 minutes, the supernatant was collected, flash frozen and stored at -20 °C until further analysis.

2.2 Antibodies
Anti-PrP monoclonal antibodies BC6, JB10 and FH10 against PrP epitopes 136-154, 216-225 and 198-207, respectively were kindly gifted by Dr Sandra McCutcheon, The Roslin Institute. For immunoblotting, antibodies were used at the following final concentrations in 0.5% (v/v) blocking reagent (Western Blocking Reagent 10 %, Roche in Tris buffered saline, pH 7.5); BC6 0.1 µg/ml, JB10 0.9 µg/ml, FH10 0.5 µg/ml, P4 (Biopharm) 0.2 µg/ml, 6H4 (Biopharm) 0.1 µg/ml, Anti- Murine α-tubulin IgG1 (Fisher Scientific) 0.01 µg/ml, horseradish-peroxidase-conjugated rabbit anti-mouse (Stratech, UK) 0.08 µg/ml. Bar224 (Bertin Pharma, SpiBio, France) which binds to the globular domain of PrP was used at a concentration of 1 µg/ml, diluted in TBS with 1% fat-free dry milk, overnight at 4°C.

2.3 Deglycosylation
Brain homogenate (10 % w/v) was denatured at 100 °C for 10 minutes and incubated with 0.125 U of peptide N-glycoside F (PNGase F kit, New England Biolabs) for 2 hours at 37 °C according to manufacturer’s instructions. Deglycosylated protein was isolated using methanol precipitation and stored at -20 °C. Before immunoblotting, the protein was pelleted by centrifugation at 10,400 g for 10 minutes. Prior to electrophoresis, samples were boiled directly in NuPAGE (Invitrogen) sample buffer supplemented with a reducing agent (Invitrogen).

2.4 SDS PAGE –Immunoblotting
Deglycosylated protein was denatured at 70 °C for 10 minutes and separated by 12% NuPAGE Bis-Tris gels (Invitrogen) or 12% Criterion gels (BioRad). Molecular markers spanning 20-220 kDa were used for size reference (MagicMarker XP Western protein Standard, Invitrogen) and electrophoresis was performed in an Xcell SureLock tank at 150 V for 1 hour using a NuPAGE kit (Invitrogen) or with the BioRad Criterion system (BioRad). Proteins were transferred onto poly(vinylidenedifluoride) membranes (Millipore or GE Healthcare) at 25 V for 1 hour, after which the membranes were washed with TBS (50 mM Tris, 150 mM NaCl, pH 7.5). The membranes were blocked using 1 % (v/v) blocking solution for 1 hour at room temperature with agitation followed by incubation with anti-PrP antibodies diluted in 0.5 % (v/v) blocking solution under the same conditions. Membranes were washed with TBST (0.1 % Tween 20 in TBS) followed by 0.5 % (v/v) blocking solution. The membranes were incubated in horseradish-peroxidase-conjugated rabbit anti-mouse (Stratech, UK) or ALP conjugated goat-anti-mouse IgG (BioRad) diluted at 1:10000 in 0.5 % (v/v)
block for 75 minutes. The membranes were washed in TBST and proteins were visualized using activated chemiluminescence (SuperSignal West Dura Extended Duration Substrate, Thermo Scientific) Lumi-Film Chemiluminescent Detection Film (Roche) or fluorescence (ALP substrate, ECF, GE Healthcare), recorded with a variable mode imager (Typhoon, GE Healthcare), when secondary antibodies labeled with ALP were used. For quantitative analysis, blots were scanned and the net intensity of manually selected protein bands representative of full length PrP²⁰, C1 and C2 were measured by use of Kodak MI software. In the case of fluorescence scans, the Image Quant Utility software (GE Healthcare) was used. The combined signal of all bands for each sample was taken as 100% and each band calculated as a percentage of the total signal. For each animal deglycosylation, immunoblotting and densitometry were performed at least twice and any further analysis was based on the average C1 value.

2.5 Production of C1 recombinant proteins
We have previously published details of full-length ovine PrP constructs [34]. Briefly, PrP open reading frame from codons 25-233 was inserted into a pTrcHis B vector and expressed in Escherichia coli strain Rosetta (DE3). To produce truncated recombinant PrP proteins representative of the C1 fragment the open reading frame between codons 115 and 234 of ovine PRNP from three genotypes (ARR/ARR, VRQ/VRQ and ARQ/ARQ) was amplified by PCR from genomic DNA with oligonucleotide primers Nde_C1F (CATCATATGTGGCAGGAGCTGCTG) and BamH1_C1R (AGTGGATCTCCTCAACTTGCCCCCCTTTG) and high fidelity Taq polymerase. The fragments were cloned into the PET 19b vector (Invitrogen) and transformed into Escherichia coli strain BL21 (DE3). The sequences of all plasmid constructs were confirmed by sequencing with T7 oligonucleotide (TAATACGACTCACTATAGG) on an AB3130 Genetic Analyzer with the BigDye® terminator v3.1 cycle sequencing kit (Applied Biosystems, USA). E.coli were cultured in terrific broth (TB) to an optical density of between 0.6-0.8 and PrP expression was induced by addition of isopropyl β-D-1-thiogalactopyranoside (IPTG) to a final concentration of 1 mM. Bacterial lysis was performed using lysozyme and inclusion bodies containing recombinant proteins were isolated and stored at -20 °C - the full details of which are described in Kirby et al (2006) [34]. Inclusion bodies were solubilised in a urea based buffer (urea 8 M, disodium hydro-phosphate 0.1 M, Tris-base 10 µM, pH8 with 2-mercaptoethanol). The full method for purification of these recombinant proteins has been previously published [35]. Briefly, recombinant proteins were first purified by nickel ion affinity chromatography, followed by desalting using a HiPrep desalting column. Proteins
were oxidised overnight and further purified by reverse phase HPLC chromatography. Purified protein was lyophilised and stored at -20 °C. All protein variants were expressed and purified on two separate occasions to ensure reproducibility of results.

2.6 Circular Dichroism
Measurements of purified protein (~0.5 mg/ml) in 50 mM sodium acetate were made using a Jasco J-710 spectropolarimeter with a path length of 0.2 mm. Readings were collected from 260 nm to 200 nm, with 20 scans repeated for each reading at a rate of 100 nm/min. Observations were made about protein secondary structure by comparison of the spectral output with proteins of known secondary structure.

2.7 Fibrillisation of recombinant C1
Fibrillisation followed a method previously published [36, 37]. In summary, lyophilised full length and C1 rPrP were reconstituted in 6 M guanidine-HCl (pH 6.0) to a concentration of 3 mg/ml. Reaction mixtures contained 2 M guanidine-HCL, 50 mM MES, 10 mM thiourea and 100 µg/ml PrP/C1. For assays with two proteins, each was added to a final concentration of 50 µg/ml. To monitor fibrillation kinetics, Thioflavin (ThT) was added (10 µM). 167 µL of reaction mix was dispensed into each well of a 96 well plate along with 3 Teflon balls (2.381 mm diameter, The Precision Plastic Ball Company). The plate was incubated with shaking, 900rpm at 37 °C in a fluorescent plate reader (Fluoroskan Ascent, Thermo Scientific). ThT fluorescence was measured every 5 minutes for 24 hours (excitation at 444 nm, emission at 485 nm). Data were analysed and lag times calculated as described in Graham et al 2010 [37]. To account for background fluorescence, fibrillation reactions were set up in the absence of PrP protein. The number of repeat fibrillation assays (n) varied for each protein variant. For analysis of fibrils by other methods, separate reactions were set-up without ThT. Post-fibrillation, these reactions were dialysed into 10 mM sodium acetate and stored at 4 °C

2.9 Maturation and PK digestion of fibrils
This protocol was followed directly from Breydo et al [36]. Proteins were visualised by gel electrophoresis (12 % NuPAGE Bis-Tris gel, Invitrogen) and silver staining.

2.8 Electron Microscopy
Formvar coated copper grids were placed onto a 50 µl drop of fibril preparation (60 µg/mL in 10 mM sodium acetate, pH 5). After 45 seconds, the grid was removed, touched to a filter paper to remove excess fluid, and then placed onto a drop of filtered 2% aqueous phosphotungstic acid for 2 minutes. Grids were then air dried before storage and examined using a Jeol 1200EX transmission electron microscope

2.9 Statistical analysis
For analysis of all brain derived PrP\textsuperscript{C}, the non-parametric Mann-Whitney U test was applied. For \textit{in-vitro} assays, the one-tailed student T-test was applied.

3. Results

3.1 The quantity of PrP\textsuperscript{C} C1-fragment relative to total PrP\textsuperscript{C} is consistent in different areas of ovine brain.

To assess quantities of the C1 fragment in specific brain areas, we measured levels of the PrP\textsuperscript{C} C1 fragment relative to full length PrP\textsuperscript{C} by semi-quantitative analysis of Western blots using anti-PrP antibodies BC6 (Figure 1) or Bar224. Cortex, cerebellum, midbrain, thalamus, hypothalamus, medulla, pons were compared for five Roslin sheep two of ARR/ARR and three of ARQ/ARQ genotypes. A representative Western blot showing relative C1 levels in different brain regions of an ARQ homozygous sheep is shown in Figure 1A. Relative C1 levels varied considerably between animals (e.g. levels of C1 in the cortex ranged from 12-48 %), however, the ranking of percentage C1 levels between areas within a single animal appeared consistent; therefore we normalised C1 levels for all brain areas against the cortex (cortex = 1) for each animal. Applying the non-parametric Mann-Whitney U-test, no significant differences in relative C1 levels were found between cortex and cerebellum (0.98 ± 0.036), thalamus (1.36 ± 0.74), hypothalamus (1.33 ± 0.84), mid brain (1.64 ± 1.36) medulla (1.09 ± 0.73) and pons (1.17 ± 0.94). These data are summarised in Figure 1B.

To confirm that C1 levels are consistent across all brains areas in animals out-with our flock, we tested cortex, cerebellum and brainstem in a further four unrelated homozygous Norwegian sheep. C1 levels in cerebellum (0.78 ± 0.23) and brainstem (0.9 ± 0.54), relative to cortex, did not differ significantly from the Roslin sheep (data not shown). Overall, we found that relative C1 levels were similar across different brain regions for individual animals but our data suggested that relative C1 levels may vary depending on \textit{PRNP} genotype. We therefore measured relative C1 levels as a function of \textit{PRNP} genotype.

3.2 The relative abundance of the PrP\textsuperscript{C} C1 fragment in ovine cortex varies with the \textit{PRNP} genotype.

We increased the number of Roslin sheep samples per genotype to a total of 11 ARR/ARR, 13 ARQ/ARQ and 5 VRQ/VRQ. These genotypes of sheep were selected because they are associated with varying susceptibility to classical scrapie. We focused on relative C1 levels in the cortex only, since we have shown that this area is representative for most of the brain. In individual animals, relative levels of the C1 fragment ranged from 14% to 77% of total PrP\textsuperscript{C}. A representative Western blot of PrP\textsuperscript{C} in the cortex of VRQ, ARQ and ARR homozygous
sheep is shown in Figure 2(A) and illustrates the differences in C1 levels between sheep of different PRNP genotype. Similar analyses were repeated for each animal and we calculated average C1 levels within genotype groups. Animals of the ARR/ARR genotype – associated with resistance to classical scrapie – had a mean relative C1 level of 52.6% (SD ± 11.9) which was significantly higher than the mean C1 level of 27.7% (SD ± 14.5) for ARQ/ARQ (p < 0.0001) and the mean C1 level of 32.6% (SD ± 8.8) for VRQ/VRQ sheep (p ≤ 0.002). These data are shown graphically in Figure 2B. Since there was no significant difference between ARQ/ARQ and VRQ/VRQ animals (from here on designated QQ_{171}, whilst ARR homozygous sheep are designated RR_{171}), which are both associated with susceptibility to classical scrapie, we combined data from these genotype groups.

It has previously been reported in several species that relative C1 levels in the brain were around 50% of total PrP \[\text{C} \] [20-22, 25], in accordance with our ARR/ARR sheep. To confirm that the unexpectedly low mean C1 level in our 18 QQ_{171} sheep (29.1% SD ± 13) was not specific to these animals, we compared them to five Norwegian QQ_{171} sheep, which showed a mean relative C1 level of 27.8% (SD ± 7.5). Analysis with two additional antibodies FH10 and JB10 replicated the pattern seen with BC6 and Bar224 making it highly unlikely that any of these results were significantly influenced by the choice of antibody. We conclude that steady state levels of the C1 fragment relative to total PrP \[\text{C} \] were 1.8 times (1.4 - 2.1, 95% confidence interval) higher in the cortex of RR_{171} homozygotes (n = 11) compared to QQ_{171} sheep (n = 23).

3.3 Sheep of the ARR/ARR genotype typically have undetectable levels of C2 fragment.

In parallel to the C1 fragment, we measured the steady state levels of the C2 fragment also by immunoblotting with antibody BC6. Across sheep of all genotypes, the mean C2 level (for C2 \(\neq 0\)) was 7.8 % (SD ± 4.8) of total PrP \[\text{C} \], but only 11 out of 28 samples had detectable C2 (defined here as \(\geq 1\)% of total PrP \[\text{C} \]). A single RR_{171} sample out of 11 showed detectable C2, whereas in QQ_{171} samples the frequency was 10 out of 18. The RR_{171} samples were five times less likely to show C2 fragment than QQ_{171} genotypes (p ≤ 0.01), but no correlation was observed between relative levels of C1 and C2 (for C2 \(\neq 0\)).

3.4 Recombinant C1 can form amyloid fibrils in-vitro.

Our data here show that all genotypes studied express some level of C1. To further investigate genotype specific effects associated with the folding pathways of the C1 fragment, we explored in-vitro fibrillisation assays with recombinant proteins. Five proteins
were expressed, two represent full-length ovine PrP<sub>25-233</sub> of sequence variants ARR and VRQ (designated rPrP<sup>ARR</sup> and rPrP<sup>VRQ</sup>), expressed in pTrcHis B/Rosetta bacterial system. The three others represent truncated ovine PrP<sub>115-235</sub> of sequence variants ARR, VRQ and ARQ, expressed in a PET-19b/BL21 (DE3) bacterial system, which are the equivalent of the C1 fragments but which incorporate an N-terminal histidine tag to aid purification (designated rC1<sup>ARR</sup>, rC1<sup>VRQ</sup>, rC1<sup>ARQ</sup>). The regions of the protein expressed are shown schematically in Figure 3A.

As expected, the recombinant C1 proteins were detected by immune-blotting with the C-terminal antibody, BC6, but no reactivity was observed towards the N-terminal antibody P4 (Figure 3B). The purity of protein was confirmed both by SDS-PAGE stained with Instant Blue (Figure 3C) and by mass spectrometry. Mass spectrometry revealed that the rC1 proteins had been correctly expressed, but also indicated small amounts of the protein that had been modified causing mass shifts of +178Da and +258Da (see supplementary Figure S1). Such modifications have been previously documented and are attributed to spontaneous alpha-N-6-phospho-gluconoylation of the His Tag [38]. These modifications were slightly more abundant in the rC1<sup>ARR</sup> variant, but since the levels of these modifications are already low and the modifications are within the octa-histidine tag, they are unlikely to impact on protein misfolding or stability. To check this we measured circular dichroism spectra (CD) for the three rC1 protein variants, which verified that all exhibited a primarily alpha-helical structure, characterized by spectra showing two minima at 222 nm and 208 nm, as shown in Figure 3D. Using CD analysis programs SELCON3 (CONTINLL) [39, 40] it was estimated that our rC1 peptides consisted of secondary structures in the following proportions: rC1<sup>ARR</sup> was 28% (28%) α-helix, 21% (22%) β-sheet and 51% (51%) random coil; rC1<sup>VRQ</sup> was 42% (38%) α-helix, 13% (18%) β-sheet and 46% (44%) random coil; rC1<sup>ARQ</sup> was 28% (34%) α-helix, 16% (17%) β-sheet and 58% (49%) random coil.

Using the five different PrP-derived recombinant proteins, fibrillisation assays were set up in the presence of thioflavin T (ThT). Fibril formation kinetics were measured over a 24-hour period and lag times were calculated as described in Graham et al. [37]. Fibrillisation experiments were repeated multiple times and typical fibrillisation curves for C1 fragments are shown in Figure 4, whilst typical curves for the full length proteins rPrP<sup>VRQ</sup> and rPrP<sup>ARR</sup> are shown in Figure 7. Fibrillisation of rPrP<sup>ARR</sup> and rPrP<sup>VRQ</sup> showed an initial lag phase (nucleation) followed by a rapid increase in fluorescence on fibril growth (elongation) as reported previously for murine full length PrP [41, 42] and human full length PrP [43].
rPrPVRQ had an average lag time of 5.57 hours (SEM +/- 0.40, n =16) while with rPrPARR the lag time was longer (p ≤ 0.002), averaging 8.32 hours (SEM +/- 0.69, n =16).

Fibrillisation of rC1VRQ (n=24) and rC1ARR (n=14) produced curves that were qualitatively similar to those of the full-length variants, with average lag times of 3.98 (SEM +/- 0.34) and 3.96 (SEM +/- 0.22) hours respectively (Figure 4). Once maximum fluorescence was reached, ThT fluorescence curves tended to fall to around 50 % of the maximum. This phenomenon of reducing fluorescence signal following fibril formation has previously been reported during the fibrillisation of human rPrP90-231. The reasons for it are yet to be fully explained, but Almstedt et al. suggested that aggregation of fibrils was the most likely cause, thereby reducing the outward facing regions available for ThT binding [43].

Compared to other rC1 proteins, rC1ARR (n=24) showed a modified fibrillisation profile comprising a significantly longer lag time (p ≤ 2.5 x 10^-6), on average 11.0 hours (SEM +/- 1.0) (Figure 4) and a reduced elongation rate after nucleation. Indeed, complete fibrillisation often did not occur within the 24-hour time frame of the experiment; in these cases the lag times could not be calculated, since this calculation depends on maximal levels of ThT fluorescence being achieved. Figure 4 shows an example of two rC1ARR reactions, one positive and one negative for complete fibrillisation. These results indicate that rC1ARR has a decreased ability to form amyloid fibrils when compared to other two rC1 variants.

To confirm that C1 proteins were producing fibrils rather than non-fibrillar, ThT-binding aggregates, we applied a fibril maturation assay originally described by Bocharova et al [44]. Fibrils composed of rPrP have a PK resistant core of10-12 kDa, but by heating PrP amyloid fibrils to 80 °C a structural change is initiated such that the PK-resistant core is extended to 16 kDa. Protease-resistant PrP can be detected SDS-PAGE and silver staining and typical results are shown in Figure 5. For the full length rPrP samples, a well-defined 16 kDa band is evident after maturation along with the characteristic 12 and 10 kDa fragments, consistent with previously reported results [36, 44] and confirming the presence of fibrils. In rC1 fibrillisation reactions along with the three bands described above, two additional PK-resistant fragments appeared after maturation, one of the similar size to the intact rC1 protein (~16.6 kDa) and one of ~13 kDa. The presence of a 13 kDa band was also evident in maturation assays of human PrP90-231 [45] and may indicate formation of disordered aggregates with increased PK resistance. The 16.6 kDa band could represent an increase in core expansion or a structural change which alters available PK digestion sites. The 16 kDa
band was less intense in rC1 fibril preparations compared to rPrP preparations, suggesting that a reduced amount of fibrils may be produced in rC1 fibrillisation reactions. To further characterise rC1 amyloid fibrils, we analysed preparations by electron microscopy, as shown in Figure 6. Fibrils were observed in all preparations and there were no clear morphological differences between fibrils from different samples, although using this technique it was not possible to quantify the number of fibrils. Nevertheless, it is clear that rC1 proteins all form fibrils in a similar manner to rPrP, but that fibrillisation of the rC1ARR protein proceeds with an extended lag time and reduced growth rate compared to rC1VRQ and rC1ARQ.

3.5 rC1ARR but not rC1VRQ or rC1ARQ inhibits fibrillisation of other PrP variants. In vivo, it is possible that the C1 fragment may interfere with misfolding of full length PrPC. Depending on their PRNP genotype, ovine cells will contain a mixture of full-length PrPC molecules and C1 fragments with identical or different protein sequences at ratios that vary from 4:1 to 1:1. To partially mimic this in vivo scenario, we performed mixed fibrillisation assays with both truncated and full length recombinant proteins at different ratios. Fibrillisation reactions were repeated multiple times and typical results for selected reaction mixtures are show in Figure 7. rPrPVRQ was mixed at a 1:1 ratio with rC1ARR (n=10), rC1ARQ (n = 6) or rC1VRQ (n = 10) and the lag time of fibrillisation was measured. There was no significant change in lag time with addition of rC1ARQ (data not shown) or rC1VRQ compared to rPrPVRQ alone. However, rC1ARR addition significantly increased the lag time of fibrillisation of rPrPVRQ from 5.57 hours (SEM ± 0.40) to 10.6 hours (SEM ± 1.12) (p = 5.2 x 10^{-5}). As observed with fibrillisation of rC1ARR, approximately 75% of the mixed assays containing rC1ARR did not result in full fibrillisation within the time frame of the experiment and therefore lag times could not be calculated in these cases. When full fibrillisation was achieved, it was evident that the elongation phase of fibrillisation of rPrPVRQ was extended and closely resembled those observed with rC1ARR alone, with a slow ThT increase after nucleation and low maximum fluorescence.

The presence of fibrils in these mixed assays was confirmed using the maturation assay, as shown in Figure 5, which showed 10 and 12 kDa bands as described for both rPrP species along with a prominent 16.6 kDa band associated with C1 fibrils alone after maturation. The 16 kDa band which would confirm the presence of full length PrP fibrils is very faint when compared to maturation and PK digestion of full length rPrP alone, indicating that full length PrP has not formed amyloid fibrils but this mixed reaction favours the production of C1 fibrils.
Reactions at a 4:1 rPrP\textsuperscript{VRQ} : rC1\textsuperscript{ARR} ratio were also performed (n = 7), in which average lag time for fibrillisation of rPrP\textsuperscript{VRQ} was increased to 6.6 hours (SEM ± 1.36). In these reactions, fibrillisation of rPrP\textsuperscript{VRQ} was inhibited to less of an extent compared to the 1:1 reactions. The rate of elongation and maximum fluorescence were lower than seen with full length protein alone, with 3 out of 7 repeats failing to fibrillise fully within the time frame of the experiment (Figure 7).

We then mixed rPrP\textsuperscript{ARR} with all three variants of C1 at a 1:1 ratio. Addition of rC1\textsuperscript{ARR} to rPrP\textsuperscript{ARR} (n=10) increased the average lag time only marginally from 8.3 (SEM ± 0.69) to 9.0 (SEM ± 0.41), but reduced the rate of fibril elongation dramatically, as illustrated in Figure 7. In contrast, addition of rC1\textsuperscript{VRQ} or rC1\textsuperscript{ARQ} (data not shown) to rPrP\textsuperscript{ARR} shortened the lag time of fibrillisation. The early ThT fluorescence increase may represent C1 fibrillisation and mask the slightly later fibrillisation of rPrP\textsuperscript{ARR} or the addition of these C1 variants may increase the rate of nucleation and elongation of rPrP\textsuperscript{ARR}. Further evidence to support the second of these hypotheses is the observation, in some mixed reactions, of two peaks of fluorescence, illustrated in the rPrP\textsuperscript{VRQ} + rC1\textsuperscript{VRQ} fibrillisation curve in Figure 7. In these cases, curves of best fit could not be produced and lag times were not calculated.

We propose that the two fluorescence maxima may represent separate fibrillisation mechanisms of the two protein variants, however, at this point we have no direct evidence of this. Across all our fibrillisation reactions, we calculated lag times, for those reactions that produced a maximum of ThT fluorescence, and averaged these. The results are shown graphically in Figure 8, from which it is clear that the rC1\textsuperscript{ARR} inhibits fibrillisation of both full length rPrP\textsuperscript{ARR} and rPrP\textsuperscript{VRQ} whilst rC1\textsuperscript{VRQ} and rC1\textsuperscript{ARQ} do not inhibit fibrillisation.

4. Discussion

The susceptibility of sheep to various natural prion diseases provides a valuable genetic model of disease resistance in which mechanisms of protein misfolding and neurodegeneration can be studied [16, 33, 46]. PRNP genetics are exploited worldwide in breeding programmes to manage the risk of scrapie outbreaks, despite a critical lack of understanding the underlying mechanisms that link PrP\textsuperscript{C} variants to disease resistance. By revealing a molecular basis of resistance, more effective measures could be put in place both to restrict new prion strains from adapting to low risk PRNP genotypes and to maintain PrP\textsuperscript{C} variants with protective potential.

A number of \textit{in vitro} studies into the molecular aspects of resistance of specific sheep genotypes to scrapie have probed the initial binding of cellular to disease-associated prion
protein prior to conversion [47, 48] or on the relative convertibility of prion protein variants [15, 34]. Others have tried to explain disease resistance by allele-specific expression profiles [47, 48], different protein conformations or different characteristics of fibrillisation and thermal stability [14]. Our results demonstrate that there are higher levels of PrP$^C$ α-cleavage fragments in sheep of the scrapie-resistant genotype ARR/ARR relative to sheep of more susceptible genotypes. The C1 fragment is produced by the action of an unknown ‘α-secretase’ and high C1 levels in ARR/ARR sheep may result from an increased proteolytic processing of PrP$^{ARR}$ compared to PrP$^{ARQ}$ or PrP$^{VRQ}$ or due to differences in the half-lives of the allotypic C1 fragments. The former may be explained by conformational differences leading to increased presentation of the cleavage site for the α-secretase or by the extended presence of PrP$^{ARR}$ within the Golgi apparatus increasing the chance of interacting with the α-secretase [49].

Setting the reason for the differences aside, the correlation between a high C1 level and scrapie resistance is supported by similar evidence from cell culture assays and transgenic mouse models. Lewis et al [32] showed that neuronal and non-neuronal cell lines with higher levels of α-cleavage show more resistance to prion infection and concluded that this cleavage is the most important factor in the selective vulnerability to prions [32]. Scrapie challenges in transgenic mice expressing only C1 protein [31] showed that C1 alone does not act as a substrate for conversion to PrP$^{Sc}$. Over-expression of C1 in the presence of full length PrP$^C$ slowed accumulation of PrP$^{Sc}$ and extended the incubation period of disease in these transgenic mice, supporting the view that C1 may inhibit some unknown pathway in the conversion of PrP$^C$ to PrP$^{Sc}$.

Along with the possible effects of C1, the corresponding product of α-cleavage, N1 has been shown to possess neuroprotective functions. For example, N1 can inhibit staurosporine induced caspase-3 activation through the p53 pathway [50] and addition of a recombinant N1 isoform in cultured cells also shows dose dependent neuroprotective effects [51]. Although we have not measured levels of N1 in our sheep brain samples, we can infer its presence from the levels of the corresponding C1 fragment and it is likely, therefore, that the increased levels of this fragment in disease-resistant animals will represent an additional benefit in protection against the neurodegeneration associated with prion disease.

Our observation that relative C2 levels in ARR/ARR sheep were significantly reduced when compared to the other genotypes is surprising, since there is little supporting evidence of an inverse correlation between α – and β-cleavage in healthy individuals. The small C2 percentage in sheep with QQ171 genotypes may represent an additional susceptibility-
enhancing factor which remains to be investigated in more detail. Truncated recombinant PrP (90-231), representing the C2 isoform, has a similar ability to aggregate and fibrillise as full length recombinant PrP [52].

Since PrPARR associated with prion resistance produces increased C1 levels in vivo and the presence of C1 protein reduces PrP conversion levels in cell culture [32] and transgenic mice [31] we investigated whether the C1 fragment could interfere with misfolding of PrP. We produced various allotypes of recombinant C1 and full length PrP and assessed their ability to form amyloid fibrils in vitro. All variants had the ability to form amyloid fibrils as confirmed by increases in ThT fluorescence during fibrillisation, the presence of a 16 kDa protease-resistant core after maturation [44] and the presence of fibrillar aggregates visualised by electron microscopy.

rC1ARR had a significantly longer lag time of fibrillisation and an altered fibril growth phase when compared to both other allotypic forms of rC1 and rPrPC. Often reactions with rC1ARR would fail to fibrillise fully within the time frame of the experiment indicating that this protein has a reduced potential to form fibrils. Similar polymorphism dependent differences have been reported with in vitro conversion of human rPrP90-231 [53-55].

When mixed with other protein variants at a ratio of 1:1, to represent a situation similar to heterozygous brain tissue, rC1ARR inhibited fibril formation whilst rC1ARQ and rC1VRQ did not. In fibrillisation reactions with a rPrP:C ratio of 0:1, inhibition was observed but the increase in lag time was less dramatic suggesting that the rPrP:rC1 ratio could be an important factor if similar inhibitory mechanisms apply in vivo.

Our findings show that rC1ARR can inhibit the initial nucleation of amyloid fibrils and slow the elongation phase of fibrillisation, but it is unclear if C1 also prevents conversion of PrPC to PrPSc. Westergard et al hypothesised that C1 acts as competitor for the binding of PrPC to PrPSc, thereby delaying the onset of disease [31]. We expand on these results by showing that rC1 can directly inhibit the formation of rPrP amyloid fibrils in the absence of PrPSc and we believe that there may be two mechanisms explaining this inhibition. The rC1 protein may bind to rPrP and prevent further interaction with other rPrP proteins thereby inhibiting the formation of an initiating nucleus. Alternatively, the rC1 protein may become incorporated into amyloid fibrils of rPrP, which would interfere with subsequent elongation of the fibrils, perhaps by inducing the formation of disordered aggregates. Since the addition of rC1ARR both increases the lag time and alters the elongation phase, both putative mechanisms may occur simultaneously.

In contrast to the findings by Westergard et al in our in vitro assays only one out of three
ovine C1 sequences had an inhibiting effect on fibril formation, indicating it is the arginine at position 171 specifically which is responsible for the effects seen. Wild type, ovine C1<sup>ARQ</sup> appeared not to affect fibril formation in vitro, whereas mouse wild type C1, which is different from sheep C1<sup>ARQ</sup> by only nine amino acid changes - has a delaying effect on development of TSE disease in vivo [31]. It remains to be established whether this indicates that the inhibition mechanism for the seeded and un-seeded nucleation or elongation is different or that the effects of C1 on prion replication are amplified in an in-vivo situation. The dominant negative inhibitory effect of C1 may be a crucial component in defining susceptibility and incubation periods in sheep of heterozygous PRNP genotypes.

Experimental scrapie in VRQ/ARR sheep can lead to incubation periods twice as long as for VRQ/VRQ sheep and longer than VRQ/ARQ [9, 56, 57]. Furthermore, in experimental BSE, incubation periods in sheep of the ARQ/ARR genotype can be four times as long as for ARQ/ARQ sheep [58] (Goldman et al unpublished). Our data indicate that the ratio of full-length rPrP<sup>C</sup> and rC1 fragment influences the strength of inhibition. Based on our results the ratio between rC1<sup>ARR</sup> and rPrP<sup>VRQ</sup> should be close to equal in VRQ/ARR sheep, whilst in ARQ/ARR sheep there should be approximately twice as much rC1<sup>ARR</sup> than rPrP<sup>ARQ</sup>, supporting the idea that C1 fragments indeed play a part in scrapie pathogenesis.

The ability of rC1<sup>ARR</sup> to inhibit conversion of rPrP makes it a possible therapeutic target in the treatment of prion disease. The possible effects of C1 protein if applied exogenously or increasing the levels of α-cleavage through manipulation of α-secretase in the in the early stages of TSE infection may slow disease progression, and warrants further investigation.

**Acknowledgements**

We would like to thank Dr Sandra McCutcheon for PrP specific monoclonal antibodies, James Foster and Roslin Institute farm staff for sampling of sheep tissues, Sonia Agarwal, Dr Charmaine Love and Paula Stewart for the optimisation of methods (all Roslin Institute) and Prof Arild Espenes, Berit Christophersen for sampling of abattoir material (Norwegian School of Veterinary Science). L.C., A.C.G., J.H., N.H. and W.G. were supported through a strategic program grant to The Roslin Institute by the Biotechnology and Biological Sciences Research Council, UK.
References


[33] N. Hunter, J.D. Foster, W. Goldmann, M.J. Stear, J. Hope, C. Bostock, Natural scrapie
ina closed flock of Cheviot sheep occurs only in specific PrP genotypes, Arch. Virol. 141 (1996) 809-824.


[48] S. McCutcheon, N. Hunter, F. Houston, Use of a new immunoassay to measure PrP Sc


Figure 1 – Measurement of C1 fragment in different brain areas of the same animal by western blot.
A: Representative Western blot of seven brain areas from an ARQ/ARQ sheep. 10% brain homogenate was treated with PNGase F and the membrane was probed with BC6 antibody. * Full length PrP\textsuperscript{C}, **C2, ***C1. These bands were measured by densitometry and C1 levels were calculated as a percentage of total PrP\textsuperscript{C} for each brain area. B: Graph comparing average levels of C1 in the cortex with six brain areas. C1 data from the six other brain areas was normalized against cortex (cortex = 1) for each animal.

Figure 2 – Comparison of relative C1 levels in the ovine cortex of sheep with varied susceptibility to scrapie.
A: Representative Western blots from each genotype tested (ARR/ARR, ARQ/ARQ and VRQ/VRQ). 10% brain homogenate was treated with PNGase F and the membrane was probed with BC6 antibody. * Full length PrP\textsuperscript{C}, **C2, ***C1. B: Graph showing the average C1 values as a percentage of total PrP as measured by densitometry. Each sample was blotted and measured a minimum of twice and the average for each animal was plotted, ARR/ARR (n= 11), ARQ/ARQ (n=13), VRQ/VRQ (n=5). An overall average for each genotype was calculated and compared statistically using the Mann-Whitney U test.

Figure 3 – Characterisation of rC1 proteins.
A: Diagrammatic representation of recombinant full length and truncated C1 protein. C1 was expressed with His-Tag to aid purification in the absence of the N-terminal region. B: Western blots of rC1 proteins, probed with monoclonal antibodies P4 (epitope 89-104) and BC6 (epitope 146-154). Lane 1: rC1\textsuperscript{ARR}; Lane 2: rC1\textsuperscript{ARQ}, Lane 3: rC1\textsuperscript{VRQ}. rC1 was not recognized by P4 but showed binding to BC6. C: All variants of recombinant protein expressed were separated by SDS-page and stained with Instant Blue to confirm purity. Lane 1: rPrP\textsuperscript{ARR}, Lane 2: rPrP\textsuperscript{VRQ}, Lane 3: rC1\textsuperscript{ARR}, Lane 4: rC1\textsuperscript{ARQ}, Lane 5: rC1\textsuperscript{VRQ}. D: Circular Dichroism analysis of three C1 variants to assess secondary structure.

Figure 4 – Kinetics of rC1 fibrilisation measured by ThT fluorescence.
Typical fibrillisation curves for each rC1 variant. rC1\textsuperscript{ARQ} and rC1\textsuperscript{VRQ} have short lag times followed by a rapid elongation phase. rC1\textsuperscript{ARR} often did not reach fibrillisation within the time frame of the experiment, described as negative for fibrillisation (rC1\textsuperscript{ARR}-ve).

**Figure 5 – Maturation and PK digestion of PrP and C1 fibrils.**

0.5 µg of recombinant protein was untreated (Lane 1), PK treated at a ratio of PrP:PK (Lane 2), or matured by heating to 80°C followed by PK treatment (Lane 3). Matured fibrils (Lane 3) show increased PK resistance relative to non-matured fibrils. rPrP reactions show a 16 kDa band after maturation which confirms the presence of fibrils, highlighted by the black box. In rC1 preparations, bands with increased PK resistance of 16.6 kDa, 16 kDa and 13 kDa are present, highlighted by the dotted box. Mixed reactions of rPrP\textsuperscript{VRQ} and rC1 have a prominent band of 16.6 kDa with a faint 16kDa band, highlighted by the dashed box.

**Figure 6 – Electron Microscopy of rC1 and rPrP fibrils.**

Fibrils were produced in the absence of ThT and dialysed into sodium acetate. Fibrils were viewed on a transmission electron microscope. Images were taken at x 5000 and x 20000. A: rC1\textsuperscript{ARR}, B: rC1\textsuperscript{ARQ}, C: rC1\textsuperscript{VRQ}, D: rPrP\textsuperscript{ARR}, E: rPrP\textsuperscript{VRQ}.

**Figure 7 – Fibrillisation of rPrP in the presence of rC1.**

Typical fibrillisation curves for mixed reactions. Addition of C1\textsuperscript{ARR} to full length PrP\textsuperscript{ARR} and PrP\textsuperscript{VRQ} at both a 1:1 and a 4:1 ratio extends the lag time of fibrillisation, slows the rate of ThT fluorescence increase and lowers maximum fluorescence. These effects are reduced in experiments at 4:1 ratio. During mixed reactions, two peaks of fluorescence were evident in some wells, illustrated in rPrP\textsuperscript{VRQ} + rC1\textsuperscript{VRQ} 1:1.

**Figure 8 – Addition of rC1\textsuperscript{ARR} increases lag time of fibrillisation of other rPrP variants.**

A) Graph showing average lag time for rPrP\textsuperscript{VRQ}: rC1\textsuperscript{ARR}, rC1\textsuperscript{ARQ}, rC1\textsuperscript{VRQ} mixed fibrillisation reactions along with the frequency of reactions failing to fibrillise within the 24 hour time frame of the experiment, calculated as a percentage of total reactions. B) As described for graph A) for rPrP\textsuperscript{ARR} : rC1\textsuperscript{ARR}, rC1\textsuperscript{ARQ}, rC1\textsuperscript{VRQ} mixed fibrillisation reactions.

Supplementary Figure S1 Mass spectrum of rC1\textsuperscript{ARR} and rC1\textsuperscript{VRQ}
Figure 1

A.

B.
Figure 2

A.

B.
Figure 3

A.

B.

C.

D.
Figure 6
Figure 7

The figure shows the relative fluorescence over time for different combinations of rPrP_{ARR} and rC1_{VRQ}. The graphs are labeled as follows:

- **rPrP_{ARR}**
- **rC1_{ARR}**
- **rPrP_{VRQ}**
- **rC1_{VRQ}**

The combinations and their corresponding graphs are:

- rPrP_{ARR} + rC1_{ARR} 1:1
- rPrP_{VRQ} + rC1_{ARR} 1:1
- rPrP_{ARR} + rC1_{ARR} 4:1
- rPrP_{VRQ} + rC1_{ARR} 4:1
- rPrP_{ARR} + rC1_{VRQ} 1:1
- rPrP_{VRQ} + rC1_{VRQ} 1:1

The x-axis represents time in hours, and the y-axis represents relative fluorescence.
Figure 8

A

![Graph A]

B

![Graph B]
Supplementary Figure S1

(A) Mass spectrum and (inset) deconvoluted mass spectrum of the expressed rC1\textsuperscript{ARR} fragment.

The spectrum consists of a single major charge envelope, labelled with ‘A’ along with the charge state of each of the species (20+, 19+ etc) and the mass/charge ratio at which each peak occurs. Deconvolution of the spectrum shows the presence of a single major species (labelled Mr(A) 16505.9) which represents the unmodified protein. The measured molecular weight of the unmodified protein (16505.9 Da) is in excellent agreement with the theoretical molecular weight calculated from the sequence of 16506.1 Da. A second species (labelled Mr(B) 16764.1) occurs at a molecular weight of 258 Da higher than the unmodified protein whilst a third (labelled Mr (C) 16684.7) is 178 Da higher in mass than the unmodified protein. These species are consistent with previously reported modifications within the His Tag region of the protein, specifically attributed to spontaneous alpha-N-6-phospho-gluconoylation.
(A) Mass spectrum and (inset) deconvoluted mass spectrum of the expressed rC1VRQ fragment.

The spectrum consists of a single major charge envelope, labelled ‘A’ along with the charge state of each of the species (20+, 19+ etc) and the mass/charge ratio at which each peak occurs. Deconvolution of the spectrum shows the presence of a single major species (labelled Mr(A) 16505.9) which represents the unmodified protein. The measured molecular weight of the unmodified protein (16505.9 Da) is in excellent agreement with the theoretical molecular weight calculated from the sequence of 16506.1 Da. A second species (labelled Mr(B) 16763.5) occurs at a molecular weight of 258 Da higher than the unmodified protein whilst a third (labelled Mr (E) 16684.1) is 178 Da higher in mass than the unmodified protein. These species are consistent with previously reported modifications within the His Tag region of the protein, specifically attributed to spontaneous alpha-N-6-phosphogluconoylation. The levels of modifications are lower for the VRQ variant than for the ARR variant, but in both cases appear to be present in less than 15% of the protein.